CSC 261/461
Database Systems

Eustrat Zhupa

October 6, 2018

UNIVERSITY of ROCHESTER

Functional Dependencies

Third Normal Form

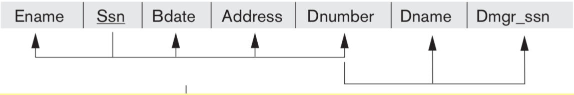
- `Transitive functional dependency`: a FD $X \rightarrow Z$ that can be derived from two FDs $X \rightarrow Y$ and $Y \rightarrow Z$
- A relation schema R is in `third normal form` (3NF) if it is in 2NF and no non-prime attribute A in R is transitively dependent on the primary key.

Functional Dependencies

Third Normal Form

(b)

EMP_DEPT



Functional Dependencies

Normal Forms

- `1st` normal form:

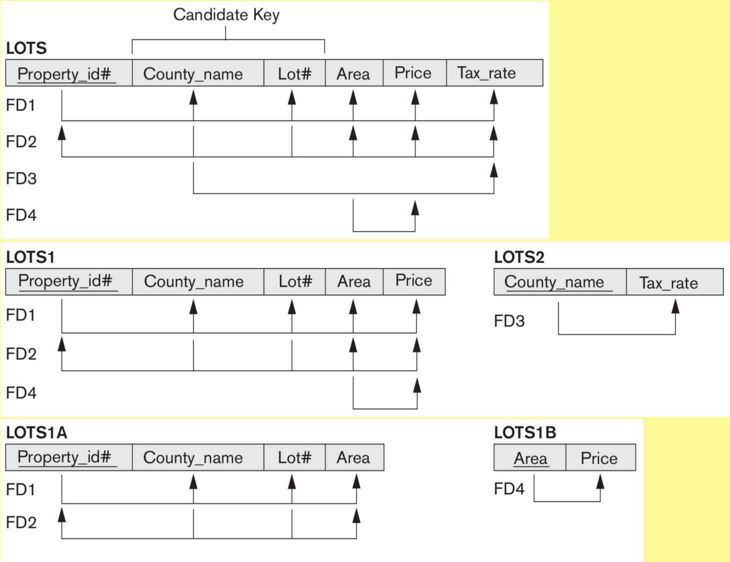
  All attributes depend on the key
- `2nd` normal form:

  All attributes depend on the whole key
- `3rd` normal form:

  All attributes depend on nothing but the key

Functional Dependencies

Normal Forms



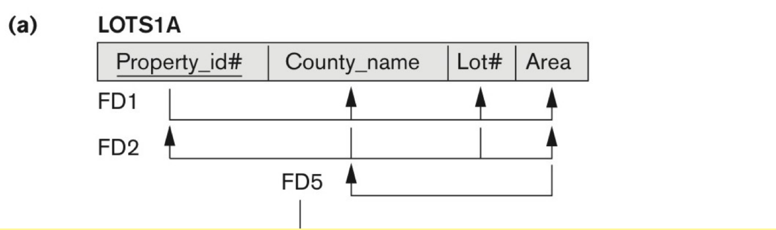
Normal Forms

Boyce-Codd Normal Form

- A relation schema R is in Boyce-Codd Normal Form (BCNF) if whenever an FD $X \to A$ holds in R, then X is a superkey of R
- Each normal form is strictly stronger than the previous one
  - Every 2NF relation is in 1NF
  - Every 3NF relation is in 2NF
  - Every BCNF relation is in 3NF
- There exist relations that are in 3NF but not in BCNF
- BCNF is considered a stronger form of 3NF
- The goal is to have each relation in BCNF (or 3NF)

Normal Forms

Boyce-Codd Normal Form



Normal Forms

Boyce-Codd Normal Form

TEACH

Student	Course	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar

Functional Dependencies

Functional Dependency

- $X \rightarrow Y$ holds if whenever two tuples have the same value for X, they must have the same value for Y
- For any two tuples t1 and t2 in any relation instance r(R): If $t1[X] = t2[X]$, then $t1[Y] = t2[Y]$
- $X \rightarrow Y$ in R specifies a constraint on all relation instances $r(R)$
- FDs are derived from the real-world constraints on the attributes

Functional Dependencies

- *Definition*: An FD $X \rightarrow Y$ is `inferred` from a set of dependencies F specified on R if $X \rightarrow Y$ holds in every legal relation state r of R; that is, whenever r satisfies all the dependencies in F, $X \rightarrow Y$ also holds in r.
- Given a set of FDs F, we can infer additional FDs that hold whenever the FDs in F hold

Functional Dependencies

Armstrong's inference rules

- IR1. (Reflexive) If $Y \subseteq X$, then $X \to Y$
- IR2. (Augmentation) If $X \to Y$, then $XZ \to YZ$
- IR3. (Transitive) If $X \to Y$ and $Y \to Z$, then $X \to Z$
- IR1, IR2, IR3 form a `sound` and `complete` set of inference rules

Functional Dependencies

Other Inference Rules

- Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$ and $X \rightarrow Z$
- Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$
- Pseudotransitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$, then $WX \rightarrow Z$
- The last three inference rules, as well as any other inference rules, can be deduced from IR1, IR2, and IR3 (completeness property)

Functional Dependencies

- Closure of a set F of FDs is the set F^+ of all FDs that can be inferred from F
- Closure of a set of attributes X with respect to F is the set X^+ of all attributes that are functionally determined by X
- X^+ can be calculated by repeatedly applying IR1, IR2, IR3 using the FDs in F.

Functional Dependencies

Algorithm for X^+

- **Input**: A set F of FDs on a relation schema R, and a set of attributes X, which is a subset of R.

$X^+ := X$
repeat
  $oldX^+ := X^+$
for each functional dependency $Y \rightarrow Z$ in F do
  if $Y \subseteq X^+$ then $X^+ := X^+ \cup Z$
until $(X^+ = oldX^+)$;

Functional Dependencies

Example

Consider the following relation schema about classes held at a university in a given academic year.

- CLASS (Classid, Course#, Instr_name, Credit_hrs, Text, Publisher, Classroom, Capacity).
- Let F, the set of functional dependencies for the above relation include:
- FD1: Sectionid $\rightarrow$ Course#, Instr_name, Credit_hrs, Text, Publisher, Classroom, Capacity
- FD2: Course# $\rightarrow$ Credit_hrs
- FD3: {Course#, Instr_name} $\rightarrow$ Text, Classroom
- FD4: Text $\rightarrow$ Publisher
- FD5: Classroom $\rightarrow$ Capacity